

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное  
образовательное учреждение высшего образования  
«Астраханский государственный университет имени В. Н. Татищева»  
(Астраханский государственный университет им. В. Н. Татищева)

СОГЛАСОВАНО  
Руководитель ОПОП  
О.Н. Выборнова  
«05» мая 2025 г.

УТВЕРЖДАЮ  
И.о. заведующего кафедрой  
информационной безопасности  
В.А. Черкасова  
«05» мая 2025 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ(МОДУЛЯ)**  
**«Вероятностно-статистические методы в анализе данных»**  
*наименование дисциплины (модуля)*

Составитель(-и)	<b>Мартьянова А.Е., доцент, к.т.н., доцент кафедры информационной безопасности</b>
Согласовано с работодателем	<b>Лазарев Н.В., инженер второй категории группы контроля безопасности объектов критической информационной инфраструктуры отдела информационной безопасности управления корпоративной защиты ООО «Газпром добыча Астрахань»</b>
Направление подготовки	<b>10.03.01 ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ</b>
Направленность (профиль) ОПОП	<b>«Организация и технологии защиты информации (в сфере информационных и коммуникационных технологий)»</b>
Квалификация (степень)	<b>бакалавр</b>
Форма обучения	<b>очная, очно-заочная</b>
Год приема (курс)	<b>2025</b>
Курс	<b>2 (по очной форме) 2 (по очно-заочной форме)</b>
Семестры	<b>3 (по очной форме) 3 (по очно-заочной форме)</b>

Астрахань, 2025

## **1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ(МОДУЛЯ)**

**1.1. Целью освоения дисциплины** «Вероятностно-статистические методы в анализе данных» является формирование у студентов современных теоретических знаний о вероятностных и статистических закономерностях, практических навыков в решении и исследовании прикладных задач теоретико-вероятностного и статистического характера, выработка у студентов теоретико-вероятностной интуиции, необходимой при решении разнообразных задач, возникающих в профессиональной деятельности.

### **1.2. Задачи освоения дисциплины (модуля):**

- приобретение студентами необходимых знаний основ вероятностно-статистических методов в анализе данных;
- овладение навыками исследования случайных величин, вычисления их основных характеристик, статистического анализа выборок;
- приобретение знаний и навыков моделирования случайных событий, обработки статистических данных, точечного и интервального оценивания параметров распределений, проверки статистических гипотез;
- формирование умения интерпретировать результаты вероятностных и статистических исследований и применять их при решении практических задач;
- применение полученных знаний, навыков и умений в последующей профессиональной деятельности.

## **2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОПОП**

**2.1. Учебная дисциплина (модуль)** «Вероятностно-статистические методы в анализе данных» относится к обязательной части плана и осваивается в 3 семестре.

**2.2. Для изучения данной учебной дисциплины (модуля) необходимы следующие знания, умения и навыки, формируемые предшествующими учебными дисциплинами:**

- Математика.
- Информатика.

В результате освоения этих дисциплин, студент должен иметь:

Знания:

- основных понятий и методов математического анализа;
- основных понятий и методов аналитической геометрии;
- основных понятий и методов линейной алгебры и теории алгебраических систем;
- основных понятий информатики.

Умения:

- использовать математические методы для решения прикладных задач;
- использовать программные и аппаратные средства персонального компьютера.

Навыки:

- владения методами количественного анализа процессов обработки, поиска и передачи информации;
- поиска информации в глобальной информационной сети Интернет и работы с офисными приложениями (текстовыми процессорами, электронными таблицами, средствами подготовки презентационных материалов, СУБД и т.п.).

**2.3. Последующие учебные дисциплины (модули), для которых необходимы знания, умения и навыки, формируемые данной учебной дисциплиной (модулем):**

- Сети и системы передачи информации;
- Методы и средства криптографической защиты информации;
- Анализ и оценка рисков;

- Анализ данных в информационной безопасности.

### 3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Процесс изучения дисциплины направлен на формирование элементов следующих компетенций в соответствии с ФГОС ВО и ОПОП ВО по данному направлению подготовки (специальности):

а) общепрофессиональных (ОПК): ОПК-3: Способен использовать необходимые математические методы для решения задач профессиональной деятельности

**Таблица 1. Декомпозиция результатов обучения**

Код компетенции	Код и наименование компетенции	Планируемые результаты обучения по дисциплине (модулю)		
		Знать (1)	Уметь (2)	Владеть (3)
ОПК-3	ОПК-3.1. Способен использовать необходимые математические методы для решения задач профессиональной деятельности	основы математики, основные математические методы.	решать стандартные профессиональные задачи с применением методов математического анализа и моделирования.	навыками математического исследования объектов профессиональной деятельности.

### 4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Общая трудоемкость дисциплины в соответствии с учебным планом составляет 3 зачетные единицы (108 часов).

Трудоемкость отдельных видов учебной работы студентов очной, очно-заочной форм обучения приведена в таблице 2.1.

**Таблица 2.1. Трудоемкость отдельных видов учебной работы по формам обучения**

Вид учебной и внеучебной работы	для очной формы обучения	для очно-заочной формы обучения
Объем дисциплины в зачетных единицах	3	3
Объем дисциплины в академических часах	108	108
Контактная работа обучающихся с преподавателем (всего), в том числе (час.):	54	36
- занятия лекционного типа, в том числе:	18	18
- практическая подготовка (если предусмотрена)	0	0
- занятия семинарского типа (семинары, практические, лабораторные), в том числе:	36	18
- практическая подготовка (если предусмотрена)		

Вид учебной и внеучебной работы	для очной формы обучения	для очно-заочной формы обучения
- консультация (предэкзаменационная) <sup>1</sup>		
- промежуточная аттестация по дисциплине <sup>2</sup>		
Самостоятельная работа обучающихся (час.)	54	125
Форма промежуточной аттестации обучающегося (зачет/экзамен), семестр(ы)	зачет – 3 семестр	зачет – 3 семестр

Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий и самостоятельной работы представлены в таблице 2.2.

**Таблица 2.2. Структура и содержание дисциплины (модуля)  
для очной формы обучения**

Раздел, тема дисциплины (модуля)	Контактная работа, час.						КР / КП	СР, час.	Итого часов	Форма текущего контроля успеваемости, форма промежуточной аттестации
	Л		ПЗ		ЛР					
	Л	в т.ч. ПП	ПЗ	в т.ч. ПП	ЛР	в т.ч. ПП				
Раздел № 1. Основы Python и анализа данных.	4				8			10	22	Лабораторная работа 1
Раздел № 2. Предобработка данных	3				6			10	19	Лабораторная работа 2. Контрольная работа 1
Раздел № 3. Исследовательский анализ данных	4				8			10	22	Лабораторная работа 3
Раздел № 4. Статистический анализ данных	3				6			10	19	Лабораторная работа 4
Раздел № 5. Теория вероятностей	4				8			14	26	Лабораторная работа 5. Вопросы к зачету
<b>Консультации</b>										
<b>Контроль промежуточной аттестации</b>										<b>Зачёт</b>
<b>ИТОГО за семестр:</b>	<b>18</b>				<b>36</b>			<b>54</b>	<b>108</b>	

<sup>1</sup> Числовые данные в данной строке соответствуют трудоемкости, указанной в учебном плане в столбце «Конс. (для гр.)»

<sup>2</sup> Числовые данные в данной строке соответствуют трудоемкости, указанной в учебном плане в столбце «КПА»

**для очно-заочной формы обучения**

Раздел, тема дисциплины (модуля)	Контактная работа, час.							СР, час.	Итого часов	Форма текущего контроля успеваемости, форма промежуточной аттестации [по семестрам]
	Л	в т.ч. ПП	ПЗ	в т.ч. ПП	ЛР	в т.ч. ПП	КР / КП			
Раздел № 1. Основы Python и анализа данных.	4				4			14	22	Лабораторная работа 1
Раздел № 2. Предобработка данных	3				3			13	19	Лабораторная работа 2. Контрольная работа 1
Раздел № 3. Исследовательский анализ данных	4				4			14	22	Лабораторная работа 3
Раздел № 4. Статистический анализ данных	3				3			13	19	Лабораторная работа 4
Раздел № 5. Теория вероятностей	4				4			18	26	Лабораторная работа 5. Вопросы к зачету
<b>Консультации</b>										
<b>Контроль промежуточной аттестации</b>										<b>Зачёт</b>
<b>ИТОГО за семестр:</b>	<b>18</b>				<b>18</b>			<b>72</b>	<b>108</b>	

*Примечание:* Л – лекция; ПЗ – практическое занятие, семинар; ЛР – лабораторная работа; ПП – практическая подготовка; КР / КП – курсовая работа / курсовой проект; КПА – контроль промежуточной аттестации; КС – консультации; СР – самостоятельная работа

[При заполнении таблиц 2.2. необходимо учесть следующее:

- заполняются таблицы только по реализуемым формам обучения;
- общий объем часов на каждую тему (раздел) для разных форм обучения должен быть одинаковым;
- практическая подготовка по видам учебных занятий распределяется разработчиком РПД по темам самостоятельно в пределах часов, выделенных в учебном плане на данную дисциплину;
- самостоятельная работа по каждой теме вычисляется как разность между общим объемом часов, выделенных на тему, и количеством часов, выделенных на сумму всех видов контактной работы;
- при подсчете консультаций необходимо учесть, что в случае наличия экзамена по дисциплине проводится одночасовая консультация; разбивать часы на консультации по разделам не нужно;
- при написании курсовой работы на контактную работу с преподавателем отводится 2 часа, объем самостоятельной работы студента на курсовую работу определяется разработчиком; разбивать часы на подготовку курсовой работы по разделам и (или) темам не нужно;
- контроль промежуточной аттестации вносится в соответствующую графу и столбец, разбивать часы на КПА по разделам не нужно.

Далее в данном пункте программы размещается матрица соотнесения разделов, тем учебной дисциплины (модуля) и формируемых в них компетенций]

**Таблица 3. Матрица соотнесения разделов, тем учебной дисциплины (модуля) и формируемых компетенций**

Раздел, тема дисциплины (модуля)	Кол-во часов	Код компетенции	Общее количество компетенций
		ОПК-3	
Раздел № 1. Основы Python и анализа данных.	22	+	1
Раздел № 2. Предобработка данных	19	+	1
Раздел № 3. Исследовательский анализ данных	22	+	1
Раздел № 4. Статистический анализ данных	19	+	1
Раздел № 5. Теория вероятностей	26	+	1
<b>ИТОГО</b>	<b>108</b>		

#### **Краткое содержание каждой темы дисциплины (модуля)**

##### **Раздел № 1. Основы Python и анализа данных.**

Введение и основы синтаксиса. Списки и циклы. Операции с таблицами. Условия и функции. Библиотека Pandas. Предобработка данных. Анализ данных и оформление результатов. Веб-оболочка Jupyter Notebook. Процесс анализа данных.

##### **Раздел № 2. Предобработка данных**

Основные принципы предобработки. Работа с пропусками. Изменение типов. Поиск дубликатов. Категоризация данных.

##### **Раздел № 3. Исследовательский анализ данных**

Введение в исследовательский анализ. Графики и выводы. Изучение срезов данных. Работа с несколькими источниками данных. Взаимосвязь данных. Валидация данных.

##### **Раздел № 4. Статистический анализ данных**

Введение в статистический анализ. Описательная статистика. Математическая статистика. Выборочные методы математической статистики. Основы теории оценивания. Основы проверки статистических гипотез. Основы регрессионного анализа.

##### **Раздел № 5. Теория вероятностей**

Случайные события. Пространство элементарных событий. Статистическая вероятность. Условные вероятности. Последовательности независимых испытаний. Случайные величины. Функции и числовые характеристики случайных величин. Основные законы распределения случайных величин. Экспоненциальное распределение и его числовые характеристики. Предельные теоремы теории вероятностей.

## **5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ПРЕПОДАВАНИЮ И ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)**

### **5.1. Указания для преподавателей по организации и проведению учебных занятий по дисциплине (модулю)**

#### **Лекционные занятия**

Основной формой реализации теоретического обучения является лекция, которая представляет собой систематическое, последовательное изложение преподавателем-лектором учебного материала теоретического характера. Цель лекции — организация целенаправленной познавательной деятельности студентов по овладению программным материалом учебной дисциплины.

Порядок подготовки лекционного занятия включает в себя выполнение следующих этапов:

- изучение требований программы дисциплины,
- определение целей и задач лекции,
- разработка плана проведения лекции,
- подбор литературы (ознакомление с методической литературой, публикациями периодической печати по темам лекционного занятия),
- отбор необходимого и достаточного по содержанию учебного материала,
- определение методов, приемов и средств поддержания интереса, внимания, стимулирования творческого мышления студентов,
- написание конспекта лекции.

Лекция должна включать следующие разделы:

- формулировку темы лекции,
- указание основных изучаемых разделов или вопросов и предполагаемых затрат времени на их изложение,
- изложение вводной части,
- изложение основной части лекции,
- краткие выводы по каждому разделу,
- Заключение,
- рекомендации литературных источников по излагаемым вопросам.

### **Лабораторные занятия**

Лабораторное занятие — целенаправленная форма организации педагогического процесса, направлена на углубление научно-теоретических знаний и овладение определенными методами работы, в процессе которых вырабатываются умения и навыки выполнения тех или иных учебных действий в данной сфере науки. Они развивают научное мышление и речь, позволяют проверить знания студентов и выступают как средства оперативной обратной связи.

Правильно организованные лабораторные занятия ориентированы на решение следующих задач:

- обобщение, систематизация, углубление, закрепление полученных в процесс самостоятельной работы теоретических знаний по дисциплине (предмету),
- формирование практических умений и навыков необходимых в будущей профессиональной деятельности, реализация единства интеллектуальной и практической деятельности,
- выработка при решении поставленных задач таких профессионально значимых качеств, как самостоятельность, ответственность, точность, творческая инициатива.

Состав заданий для лабораторных занятий должен быть спланирован с расчетом, чтобы за отведенное время они могли быть качественно выполнены большинством учащихся.

Лабораторные занятия должны быть так организованы, чтобы студенты ощущали нарастание сложности выполнения заданий, испытывали бы положительные эмоции от переживания собственного успеха в обучении и овладении навыками правильных и точных решений.

### **Самостоятельная работа**

Самостоятельная работа — это вид учебной деятельности, которую студент совершает в установленное время и в установленном объеме индивидуально или в группе, без непосредственной помощи преподавателя (но при его контроле), руководствуясь

сформулированными ранее представлениями о порядке и правильности выполнения действий.

В учебном процессе образовательного учреждения выделяются два вида самостоятельной работы:

- 1) аудиторная — выполняется на учебных занятиях, под непосредственным руководством преподавателя и по его заданию (выполнение самостоятельных работ; выполнение контрольных и лабораторных работ; решение задач),
- 2) внеаудиторная — выполняется по заданию преподавателя, но без его непосредственного участия (подготовка к аудиторным занятиям; изучение учебного материала, вынесенного на самостоятельную проработку; выполнение домашних заданий разнообразного характера; выполнение индивидуальных заданий, направленных на развитие у студентов самостоятельности и инициативы; подготовка к контрольной работе). Внеаудиторные самостоятельные работы представляют собой логическое продолжение аудиторных занятий, проводятся по заданию преподавателя, который инструктирует студентов и устанавливает сроки выполнения задания.

При подготовке к лекционным занятиям необходимо воспользоваться учебно-методической литературой (основной) из п.8.

При подготовке к лабораторным занятиям необходимо воспользоваться учебно-методической литературой (дополнительной) из п.8.

## **5.2. Указания для обучающихся по освоению дисциплины (модулю)**

### **Лекционные занятия**

Лекция — основной вид обучения в вузе. В лекции излагаются основные положения теории, ее понятия и законы, приводятся факты, показывающие связь теории с практикой. Накануне лекции необходимо повторить содержание предыдущей лекции (а также теорию по изучаемой теме в учебниках), а затем посмотреть тему очередной лекции по программе (по плану лекций).

Полезно вести записи (конспекты) лекций: для непонятных вопросов оставлять место при работе над темой лекции с учебными пособиями.

Записи лекций следует вести в отдельной тетради, оставляя место для дополнений во время самостоятельной работы.

При конспектировании лекций выделяйте главы и разделы, параграфы, подчеркивайте основное.

### **Лабораторные занятия**

Лабораторное занятие — наиболее активный вид учебных занятий в вузе. Он предполагает самостоятельную работу над учебными пособиями, основной литературой, открытыми источниками информации.

К каждому лабораторному занятию нужно готовиться. Подготовку следует начинать с повторения теории (по учебному пособию). После этого нужно решать задачи из предложенного домашнего задания.

### **Самостоятельная работа**

Самостоятельность в учебной работе способствует развитию заинтересованности студента в изучаемом материале, вырабатывает у него умение и потребность самостоятельно получать знания, что весьма важно для специалиста с высшим образованием.

Самостоятельная работа студентов представлена в следующих формах:

- работа с учебной литературой и конспектом лекций с целью подготовки к лабораторным занятиям, составление конспектов тем, выносимых на самостоятельную проработку,
- систематическое выполнение домашних работ.

Во время самостоятельной работы необходимо воспользоваться учебно-методической литературой из п.8 (основной), (дополнительной), Интернет-ресурсами.

**Таблица 4. Содержание самостоятельной работы обучающихся  
для очной формы обучения**

Вопросы, выносимые на самостоятельное изучение	Кол-во часов	Формы работы
Раздел № 1. Основы Python и анализа данных. Подготовка отчета к лабораторной работе 1.	10	Внеаудиторная, изучение учебных пособий
Раздел № 2. Предобработка данных. Подготовка к лабораторной работе 2. Подготовка отчета к контрольной работе 1.	10	Внеаудиторная, изучение учебных пособий
Раздел № 3. Исследовательский анализ данных. Подготовка отчета к лабораторной работе 3.	10	Внеаудиторная, изучение учебных пособий
Раздел № 4. Статистический анализ данных. Подготовка отчета к лабораторной работе 4.	10	Внеаудиторная, изучение учебных пособий
Раздел № 5. Теория вероятностей. Подготовка отчета к лабораторной работе 5. Подготовка по вопросам к зачету.	14	Внеаудиторная, изучение учебных пособий

**для очно-заочной формы обучения**

Вопросы, выносимые на самостоятельное изучение	Кол-во часов	Формы работы
Раздел № 1. Основы Python и анализа данных. Подготовка отчета к лабораторной работе 1.	10	Внеаудиторная, изучение учебных пособий
Раздел № 2. Предобработка данных. Подготовка отчета к лабораторной работе 2. Подготовка к контрольной работе 1.	10	Внеаудиторная, изучение учебных пособий
Раздел № 3. Исследовательский анализ данных. Подготовка отчета к лабораторной работе 3.	10	Внеаудиторная, изучение учебных пособий
Раздел № 4. Статистический анализ данных. Подготовка отчета к лабораторной работе 4.	10	Внеаудиторная, изучение учебных пособий
Раздел № 5. Теория вероятностей. Подготовка отчета к лабораторной работе 5. Подготовка по вопросам к зачету.	14	Внеаудиторная, изучение учебных пособий

**5.3. Виды и формы письменных работ, предусмотренных при освоении дисциплины, выполняемые обучающимися самостоятельно**  
Не предусмотрено.

## **6. ОБРАЗОВАТЕЛЬНЫЕ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ**

При реализации различных видов учебной работы по дисциплине могут

использоваться электронное обучение и дистанционные образовательные технологии.

### 6.1. Образовательные технологии

Учебные занятия по дисциплине могут проводиться с применением информационно-телекоммуникационных сетей при опосредованном (на расстоянии) интерактивном взаимодействии обучающихся и преподавателя в режимах on-line в формах: видеолекций, лекций-презентаций, видеоконференции, собеседования в режиме чат, форума, чата, выполнения виртуальных практических и/или лабораторных работ и др.

Максимальный объем занятий обучающегося с применением электронных образовательных технологий не должен превышать 25%.

**Таблица 5. Образовательные технологии, используемые при реализации учебных занятий**

Раздел, тема дисциплины (модуля)	Форма учебного занятия		
	Лекция	Практическое занятие, семинар	Лабораторная работа
Раздел № 1. Основы Python и анализа данных.	Обзорная лекция	Не предусмотрено	выполнение лабораторной работы
Раздел № 2. Предобработка данных	Лекция-презентация	Не предусмотрено	выполнение лабораторной работы
Раздел № 3. Исследовательских анализ данных	Лекция-презентация	Не предусмотрено	выполнение лабораторной работы
Раздел № 4. Статистический анализ данных	Лекция-презентация	Не предусмотрено	выполнение лабораторной работы
Раздел № 5. Теория вероятностей	Лекция-презентация	Не предусмотрено	выполнение лабораторной работы

### 6.2. Информационные технологии

Название информационной технологии	Темы, разделы дисциплины	Краткое описание применяемой технологии
Использование возможностей Интернета в учебном процессе	1 - 5	Проведение входного, текущего и рейтингового контроля знаний учащихся (в системах дистанционного обучения)
Использование возможностей электронной почты преподавателя	1 - 5	Подготовка к защите отчетов по лабораторным работам
Использование средств представления учебной информации	1 - 5	Использование мультимедийной презентации

При реализации различных видов учебной и внеучебной работы используются следующие информационные технологии:

- - использование возможностей Интернета в учебном процессе (использование информационного сайта преподавателя (рассылка заданий, предоставление

выполненных работ, ответы на вопросы, ознакомление учащихся с оценками и т.д.);

- - использование электронных учебников и различных сайтов (например, электронные библиотеки, журналы и т.д.) как источников информации;
- - использование возможностей электронной почты преподавателя;
- - использование средств представления учебной информации (электронных учебных пособий и практикумов, применение новых технологий для проведения очных (традиционных) лекций и семинаров с использованием презентаций и т.д.);
- - использование интегрированных образовательных сред, где главной составляющей являются не только применяемые технологии, но и содержательная часть, т.е. информационные ресурсы (доступ к мировым информационным ресурсам, на базе которых строится учебный процесс);
- - использование виртуальной обучающей среды (или системы управления обучением LMS Moodle «Электронное образование») или иных информационных систем, сервисов и мессенджеров.

### 6.3. Программное обеспечение, современные профессиональные базы данных и информационные справочные системы

#### 6.3.1. Программное обеспечение

Наименование программного обеспечения	Назначение
Adobe Reader	Программа для просмотра электронных документов
Платформа дистанционного обучения LMS Moodle	Виртуальная обучающая среда
Google Chrome	Браузер
Microsoft Office 2013, Microsoft Office Project 2013, Microsoft Office Visio 2013	Офисная программа
7-zip	Архиватор
Microsoft Windows 7 Professional	Операционная система
Kaspersky Endpoint Security	Средство антивирусной защиты

#### 6.3.2. Современные профессиональные базы данных и информационные справочные системы

1. Электронный каталог Научной библиотеки АГУ на базе MARK SQL НПО «Информ-систем»: <https://library.asu.edu.ru>.
2. Электронный каталог «Научные журналы АГУ»: <http://journal.asu.edu.ru/>.
3. Универсальная справочно-информационная полнотекстовая база данных периодических изданий ООО «ИВИС»: <http://dlib.eastview.com/>
4. Электронно-библиотечная система elibrary. <http://elibrary.ru>
5. Справочная правовая система КонсультантПлюс: <http://www.consultant.ru>
6. Информационно-правовое обеспечение «Система ГАРАНТ»: <http://garant-astrakhan.ru>

## 7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

### 7.1. Паспорт фонда оценочных средств

При проведении текущего контроля и промежуточной аттестации по дисциплине (модулю) «Вероятностно-статистические методы в анализе данных» проверяется сформированность у обучающихся компетенций, указанных в разделе 3 настоящей программы. Этапность формирования данных компетенций в процессе освоения образовательной программы определяется последовательным освоением дисциплин (модулей) и прохождением практик, а в процессе освоения дисциплины (модуля) – последовательным достижением результатов освоения содержательно связанных между собой разделов, тем.

**Таблица 6. Соответствие разделов, тем дисциплины (модуля), результатов обучения по дисциплине (модулю) и оценочных средств**

Контролируемые раздел, тема дисциплины (модуля)	Код контролируемой компетенции	Наименование оценочного средства
Раздел № 1. Основы Python и анализа данных.	ОПК-3	Лабораторная работа 1.
Раздел № 2. Предобработка данных	ОПК-3	Лабораторная работа 2. Контрольная работа 1
Раздел № 3. Исследовательских анализ данных	ОПК-3	Лабораторная работа 3.
Раздел № 4. Статистический анализ данных	ОПК-3	Лабораторная работа 4.
Раздел № 5. Теория вероятностей	ОПК-3	Лабораторная работа 5. Вопросы к зачету

### 7.2. Описание показателей и критериев оценивания компетенций, описание шкал оценивания

При решении комплексной ситуационной задачи можно использовать следующие критерии оценки:

**Таблица 7. Показатели оценивания результатов обучения в виде знаний**

Шкала оценивания	Критерии оценивания
5 «отлично»	демонстрирует глубокое знание теоретического материала, умение обоснованно излагать свои мысли по обсуждаемым вопросам, способность полно, правильно и аргументированно отвечать на вопросы, приводить примеры
4 «хорошо»	демонстрирует знание теоретического материала, его последовательное изложение, способность приводить примеры, допускает единичные ошибки, исправляемые после замечания преподавателя
3 «удовлетворительно»	демонстрирует неполное, фрагментарное знание теоретического материала, требующее наводящих вопросов преподавателя, допускает существенные ошибки в его изложении, затрудняется в приведении примеров и формулировке выводов
2 «неудовлетворительно»	демонстрирует существенные пробелы в знании теоретического материала, не способен его изложить и ответить на наводящие вопросы преподавателя, не может привести примеры

**Таблица 8. Показатели оценивания результатов обучения в виде умений и владений**

Шкала оценивания	Критерии оценивания
5 «отлично»	демонстрирует способность применять знание теоретического материала при выполнении заданий, последовательно и правильно выполняет задания, умеет обоснованно излагать свои мысли и делать необходимые выводы
4 «хорошо»	демонстрирует способность применять знание теоретического материала при выполнении заданий, последовательно и правильно выполняет задания, умеет обоснованно излагать свои мысли и делать необходимые выводы, допускает единичные ошибки, исправляемые после замечания преподавателя
3 «удовлетворительно»	демонстрирует отдельные, несистематизированные навыки, испытывает затруднения и допускает ошибки при выполнении заданий, выполняет задание по подсказке преподавателя, затрудняется в формулировке выводов
2 «неудовлетворительно»	не способен правильно выполнить задания

### 7.3. Контрольные задания или иные материалы, необходимые для оценки результатов обучения по дисциплине (модулю)

#### Раздел № 1. Основы Python и анализа данных.

##### Лабораторная работа 1. Библиотека NumPy

Цель занятия:

Получение представления о функциональности, доступных методах и объектах библиотеки NumPy. Изучение основных принципов практической работы с ними.

##### 1. Создание массивов

В NumPy существует много способов создать массив. Один из наиболее простых – создать массив из обычных списков или кортежей Python, используя функцию `numpy.array()` (обратите внимание: `array()` – это метод, создающий объект типа `ndarray`):

```
In [1]: import numpy as np

a = np.array([1, 2, 3])
print(a)
print(type(a)) #вывод типа переменной a

[1 2 3]
<class 'numpy.ndarray'>
```

##### 2. Базовые операции

Математические операции над массивами выполняются поэлементно. Ниже представлены примеры математических операций между двумя массивами.

```
In [19]: import numpy as np

a = np.array([20, 30, 40, 50])
b = np.arange(4)

print ('a=',a)
print ('b=',b)

print ("n1", a + b)
print ("2", a - b)
print ("3", a * b)
print ("4", a / b) # При делении на 0 возвращается inf (бесконечность)
print ("5", a ** b) # Возведение в степень
print ("6", a % b) # При взятии остатка от деления на 0 возвращается
```

### 3. Индексы, срезы, итерации

Одномерные массивы осуществляют операции индексирования, срезов и итераций очень схожим образом с обычными списками и другими последовательностями Python.

```
In [28]: a = np.arange(10) ** 3
print(np.arange(10))
print(a)

[0 1 2 3 4 5 6 7 8 9]
[ 0  1  8 27 64 125 216 343 512 729]
```

```
In [29]: a = np.arange(10) ** 3
print ("1)", a)
print ("2)", a[1])
print("3)", a[3:7])

a[3:7] = 8
print ("4)", a)

print("5)", a[::-1])

for i in a:
    print(i ** (1/3))
```

### 4. Манипуляции с формой

У массива есть форма `shape`, определяемая числом элементов вдоль каждой оси:

```
In [49]: a = np.array([[0, 1, 2], [10, 12, 13]], [[100, 101, 102], [110, 112, 113]])
print ('a=',a)
print (a.shape) #вывод "формы", т.е. размерности

a= [[[ 0  1  2]
      [10 12 13]]

      [[100 101 102]
      [110 112 113]]]
(2, 2, 3)
```

## 5. Объединение массивов

Несколько массивов могут быть объединены вместе вдоль разных осей с помощью методов `hstack()` и `vstack()`. `hstack()` объединяет массивы по первым осям, `vstack()` – по последним.

## 6. Разбиение массива

Используя `hsplit()` вы можете разбить массив вдоль горизонтальной оси, указав либо число возвращаемых массивов одинаковой формы, либо номера столбцов, после которых массив разрезается «ножницами». Функция `vsplit()` разбивает массив вдоль вертикальной оси, а `array_split()` позволяет указать оси, вдоль которых произойдет разбиение.

Задания для лабораторной работы № 1:

1. Создание массива, заполненного нулями – приведите пример кода.
2. Математические операции между массивами. Представьте код, где массив со значениями от 1 до 9 умножается на константу 2. Должна получиться таблица умножения для числа 2.
3. Слайсы (обрезка массива) – предоставите код для четырех случаев:
  1. вывести все элементы, кроме первых трех;
  2. вывести все элементы, кроме последних трех;
  3. вывести все элементы, кроме первых и последних трех;
  4. слайсы с двумерными массивами.

## Раздел № 2. Предобработка данных

### Лабораторная работа 2. Библиотека Pandas

Цель занятия:

Получение представления о функциональности, доступных методах и объектах библиотеки `pandas`. Изучение основных принципов практической работы с ними

#### 1. Создание Series

##### Структура данных Series

Конструктор класса `Series` выглядит следующим образом:

```
pandas.Series(data=None, index=None, dtype=None, name=None, copy=False,
fastpath=False)
```

```
In [2]: s1 = pd.Series([1, 2, 3, 4, 5])
print(s1)

0    1
1    2
2    3
3    4
4    5
dtype: int64
```

## 2. Создание DataFrame

### Структура данных DataFrame

Конструктор класса DataFrame выглядит так:

```
pandas.DataFrame(data=None, index=None, columns=None, dtype=None, copy=False)
```

```
In [13]: nda1 = np.array([[1, 2, 3], [10, 20, 30]])
df4 = pd.DataFrame(nda1)
print(df4)

   0  1  2
0  1  2  3
1 10 20 30
```

## 3. Доступ к данным в структурах pandas

Основные методы для работы с объектами DataFrame представлены в таблице ниже:

Операция	Синтаксис	Возвращаемый результат
Выбор столбца	df[col]	Series
Выбор строки по метке	df.loc[label]	Series
Выбор строки по индексу	df.iloc[loc]	Series
Слайс по строкам	df[0:4]	DataFrame
Выбор строк, отвечающих условию	df[bool_vec]	DataFrame

Для доступа к данным можно использовать атрибуты структур, в качестве которых выступают метки.

## 4. Добавление элементов в структуры

Увеличение размера структуры – т.е. добавление новых, дополнительных, элементов – это довольно распространенная задача. Добавление нового элемента в Series показано ниже.

```
In [21]: s = pd.Series([10, 20, 30, 40, 50], ['a', 'b', 'c', 'd', 'e'])
print ('0)Исходный Series s\n', s)
s['f'] = 60
print ('1) Добавили s[\f'] = 60\n', s)
```

## 5. Pandas и отсутствующие данные

Метод `isnull()` из библиотеки `pandas`, который позволяет быстро определить наличие элементов `NaN` в структурах.

Подробную информацию об объекте (датафрейме), для этого можно воспользоваться методом `info()`.

## 6. Замена отсутствующих данных

Отсутствующие данные объектов можно заменить на конкретные числовые значения, для этого можно использовать метод `fillna()`.

Не всегда целесообразно заменять отсутствующие данные нулями. Иногда недостающие данные можно заполнить средним значением по столбцу:

## 7. Удаление объектов/столбцов с отсутствующими данными

Довольно часто используемый подход при работе с отсутствующими данными – это удаление записей (строк) или полей (столбцов), в которых встречаются пропуски. Для того, чтобы удалить все объекты, которые содержат значения `NaN` воспользуйтесь методом `dropna()` без аргументов.

## Задания для лабораторной работы № 2:

1. Приведите пример `Series` и попробуйте: выводить на экран, добавлять элементы, изменять значения.
2. Повторите предыдущий пункт для объекта `DataFrame`: выводить на экран, добавлять элементы, изменять значения.
3. Открыть при помощи `pandas` созданный в процессе выполнения работы `dataframe.csv` файл с пропусками или некорректными значениями и исправить их.

## Контрольная работа 1

1. Перечислите наиболее важные атрибуты объектов `ndarray`
2. Базовая функциональность `NumPy` – опишите основные доступные функции.
3. Настройка печати массивов «под свои нужды».
4. Опишите применение унарных операций к массивам, приведите примеры.
5. Копии и представления при работе с массивами.
6. Сохранение массива в файл и чтение из файла.
7. Сравнение библиотеки `Pandas` с функциональностью библиотеки `NumPy`.
8. Опишите Конструктор класса `Series`.
9. Опишите Конструктор класса `DataFrame`.
10. Опишите сохранение и чтение объектов из файла.

## Раздел № 3. Исследовательских анализ данных

## Лабораторная работа 3. Статистика в NumPy и в Statistics. Введение в Scipy.Stats

Цель занятия:

Получение представления о функциональности, доступных методах и объектах статистики библиотек/модулей NumPy, Statistics и Scipy.Stats. Изучение основных принципов практической работы с ними.

### 1. Статистика в NumPy

#### Максимум и минимум массива

Определение минимального и максимального значения элементов массива вдоль указанной оси при помощи методов `np.min()` и `np.max()`.

#### Среднее арифметическое

Среднее арифметическое (или обычно говорят кратко – «среднее») – это сумма элементов, деленная на их сумму. Метод `np.mean()` вычисляет среднее значение, складывая все элементы массивов, а затем делит его на количество элементов. Мы также можем указать ось, по которой может быть вычислено среднее значение.

#### Медиана

Медиана (с лат. «середина») – середина множества чисел. Формула отличается для множеств с четным и нечетным числом элементов. Метод `np.median()` все отсортирует сам и найдет медиану.

#### Среднеквадратичное (стандартное) отклонение

Среднеквадратичное отклонение (формула 4, так же называется стандартным отклонением) – определяется как квадратный корень из дисперсии случайной величины. При помощи метода `np.std()` вычисляется стандартное отклонение («std» – от англ. standart, стандартное).

#### Среднее взвешенное

Функция `np.average()` определяет средневзвешенное значение, и может использоваться для многомерных массивов. Средневзвешенное значение рассчитывается путем умножения компонента на его вес, веса указываются отдельно. Если веса не указаны, результат будет таким же, как и среднее значение (среднее арифметическое).

#### Процентили

Процентиль (персентиль или персентиль) – это такое значение  $X_p$  из массива (или в терминологии статистики это называется совокупностью значений), что значения  $p$ -й части совокупности меньше или равны этому значению  $X_p$ .

#### Полный размах

Функция `np.ptp()` полезна для определения размаха значений вдоль оси.

### 2. Некоторые статистические функции из statistics

#### Гармоническое среднее

Гармоническое среднее является обратным средним значением обратных величин всех элементов в наборе данных

## Среднее геометрическое

Геометрическое среднее является  $n$ -й корнем произведения всех элементов  $x_i$  в наборе данных  $X$ .

### 3. Статистика при помощи модуля `Scipy.Stats`

Этот модуль содержит большое количество распределений вероятностей, сводную и частотную статистику, корреляционные функции и статистические тесты, маскированную статистику, ядерную оценку плотности, функциональность квази-Монте-Карло и многое другое.

Все статистические функции и доступные случайные величины находятся в суб-пакете `scipy.stats`.

### Кумулятивная функция распределения (CDF).

Рассмотрим кумулятивную функцию распределения (CDF, с англ. Cumulative Distribution Function). Она показывает вероятность того, что целевое значение меньше указанного значения либо равно ему.

### Функция получения значения процентиля (PPF)

PPF (расшифровывается с англ. как percent point function) – функция, используемая для получения значения, соответствующему определенному процентилю.

Задания для лабораторной работы № 3:

1. Примеры применения среднеарифметического, медианы, среднегеометрического и гармонического среднего, процентилей.
2. Представьте код, который строит графики кумулятивной функции распределения (`cdf`), функции получения процентилей (`ppf`) и функции плотности вероятности (`pdf`) для сгенерированного при помощи NumPy набора дискретных значений от -3 до 3 с дискретностью 0.1.
3. Проведите эксперименты с библиотекой для построения графиков `matplotlib`. Попробуйте добавить `label` для графиков, `xlabel/ylabel` для осей, `title` и для всего графика целиком. На следующем занятии будет представлено решение этих задач.

## Раздел № 4. Статистический анализ данных

### Лабораторная работа № 4

Цель занятия:

Изучение функциональности, доступных методов и объектов статистики библиотеки `Scipy` – генерация случайных чисел и параметры формы распределения. Исследование принадлежности ряда значений к конкретному закону распределения (на основе параметров формы, куртозиса и др.). Построение диаграмм размах. Рассмотрение примеров описательной статистики из реального мира ( $z$ -статистика и  $p$ -значение).

### Генерация случайных чисел

Генерация случайных чисел (рядов величин случайного распределения) зависит от генераторов из модуля `numpy.random`.

Генерация по нормальному закону распределения. Генерация 5 случайных величин, распределенных по нормальному закону распределения, используя вызов `norm.rvs(size=5, random_state=rng)`.

Кроме генерации рядов с нормальным распределением, можно генерировать наборы случайных величин, с гамма распределением, или с любым другим типом распределения. Полный список которых можно найти в официальной документации `scipy` или `numpy`.

Принцип такой же (как и при рассмотренном ранее вызове `norm.rvs()`), только может отличаться набор передаваемых параметров.

### Куртозис и сдвиг (моменты, и прочие понятия статистики)

В `scipy` доступны функции и методы для получения и многих других, кроме уже рассмотренных, параметров статистики. Представлено вычисление куртозиса (острота пика в середине) и перекоса (значений относительно медианы).

### Диаграмма размаха

Диаграммы размаха в удобной форме показывают медиану (или, если нужно, среднее), нижний и верхний квартили, минимальное и максимальное значение выборки и выбросы. Расстояния между различными частями ящика так же позволяют определить степень разброса (дисперсии) и асимметрии данных.

### Примеры статистики из жизни

Пример с временем доставки. Рассмотрим реальный пример из жизни: статистика времени доставки пиццы.

В течение года мы заказываем пиццу на дом, при этом мы каждый раз смотрим на настенные стрелочные часы, отмечая время, которое проходит между заказом и доставкой, целым количеством минут. Тогда накопленные за год данные о доставке пиццы могли бы выглядеть так:

```
norm_rv = stats.norm(
    loc=30, scale=5
) # генерируем норм расп значения возле 30 плюс минус 5
samples = np.trunc(norm_rv.rvs(365))
```

Теперь выясним среднее время доставки пиццы и его среднеквадратическое отклонение:

```
samples.mean(), samples.std()
```

Можно сказать, что время доставки пиццы занимает около  $30 \pm 5$  минут. Посмотрим на то, как распределены данные:

```
sns.histplot(x=samples, discrete=True);
```

Вычислить на какую величину время доставки отличается от среднего времени доставки, Интерпретировать данный результат, используя Z-значение. Можно воспользоваться графиком:

```
fig, ax = plt.subplots()
x = np.linspace(norm_rv.ppf(0.001), norm_rv.ppf(0.999), 200)
ax.vlines(40.5, 0, 0.1, color="k", lw=2)
sns.lineplot(x=x, y=norm_rv.pdf(x), color="r", lw=3)
sns.histplot(x=samples, stat="probability", discrete=True);
```

Задания для лабораторной работы № 4:

1. Определить к какому конкретно типу распределения относится время доставки пиццы из множества всевозможных?
2. Постройте эмпирические функции распределения (относительные и накопленные частоты) для роста (в см) группы из 20 мужчин: 181, 169, 178, 178, 171, 179, 172, 181, 179, 168, 174, 167, 169, 171, 181, 181, 183, 172, 176, 181.
3. Построить эмпирические функции распределения (относительные и накопленные частоты) успеваемости в группе из 20 студентов: 4, 4, 5, 3, 4, 5, 4, 5, 3, 5, 3, 3, 5, 4, 5, 4, 3, 5, 3, 5.

## Раздел № 5. Теория вероятностей

### Лабораторная работа 5

Цель занятия:

Получение представления о вычислении интервальных оценок и доверительной вероятности, а также для проверки для проверки статистических гипотез с использованием функций `scipy.stats`. Изучение основных принципов практической работы с ними для реализации математических алгоритмов.

#### 1) Вычисление интервальных оценок и доверительной вероятности

В математической статистике **интервальной оценкой** называется результат использования выборки для вычисления интервала возможных значений неизвестного параметра, оценку которого нужно построить. Следует отличать от точечной оценки, которая даёт лишь одно значение. Самым распространенным видом интервальных оценок являются доверительные интервалы.

**Доверительная вероятность** - вероятность того, что доверительный интервал накроет неизвестное истинное значение параметра, оцениваемого по выборочным данным.

#### 2) Проверка статистических гипотез

**Статистическая гипотеза** – всякое высказывание о генеральной совокупности, которое проверяется (подтверждается или опровергается) по выборке из генеральной совокупности. Предположение о виде распределения и свойствах случайной величины, которое можно подтвердить или опровергнуть применением статистических методов к данным выборки. Гипотеза принимается или не принимается на основе анализа выборки из генеральной совокупности, следовательно может быть ошибочным как принятие так и отвержение гипотезы:

- если нулевая нулевая гипотеза отвергается, когда на самом деле она верна, то совершается ошибка **первого рода** (ложно-отрицательный ответ).
- если нулевая нулевая гипотеза принимается, когда на самом деле она не верна, то совершается ошибка **второго рода** (ложно-положительный ответ).

Для проверки гипотезы используют специально-подобранную случайную величину, которая называется статистический критерий. Наблюдаемым значением критерия называют значение критерия, вычисленное по выборке. **Критическая область** - это совокупность значений критерия, при которых нулевая гипотеза отвергается.

**Критическими точками** называют точки, отделяющие критическую область от областей принятия гипотезы.

Функции пакета `scipy.stats` для проверки гипотез обычно возвращают значение критерия и  $p$ -значение, которое сравнивается с уровнем значимости. Если  $p$ -значение больше уровня значимости, то нет основания отвергнуть нулевую гипотезу.

### Задания для лабораторной работы № 5:

Решить задачи, выданные преподавателем:

- 1) Оценка математического ожидания нормального распределения при известном  $\sigma$
- 2) Оценка математического ожидания нормального распределения при неизвестном  $\sigma$
- 3) Сравнение двух средних нормальных генеральных совокупностей, дисперсии которых известны (независимые выборки)
- 4) Сравнение выборочной средней с гипотетической генеральной средней нормальной совокупности
  - А) Дисперсия генеральной совокупности известна
  - Б) Дисперсия генеральной совокупности неизвестна
- 5) Сравнение двух дисперсий нормальных генеральных совокупностей
- 6) Сравнение выборочной (исправленной) дисперсии с гипотетической генеральной дисперсией нормальной совокупности

### Перечень вопросов к зачету

1. Основы Python и анализа данных.
2. Введение и основы синтаксиса.
3. Списки и циклы.
4. Операции с таблицами.
5. Условия и функции.
6. Библиотека NumPy.
7. Библиотека Pandas.
8. Веб-оболочка Jupyter Notebook.
9. Анализ данных и оформление результатов.
10. Процесс анализа данных.
11. Предобработка данных. Основные принципы предобработки.
12. Предобработка данных. Работа с пропусками.
13. Предобработка данных. Изменение типов.
14. Предобработка данных. Поиск дубликатов.
15. Предобработка данных. Категоризация данных.
16. Исследовательских анализ данных. Графики и выводы.
17. Исследовательских анализ данных. Изучение срезов данных.
18. Исследовательских анализ данных. Работа с несколькими источниками данных.
19. Исследовательских анализ данных. Взаимосвязь данных.
20. Исследовательских анализ данных. Валидация данных.
21. Статистический анализ данных. Описательная статистика.
22. Математическая статистика.
23. Выборочные методы математической статистики.
24. Основы теории оценивания.
25. Основы проверки статистических гипотез.
26. Основы регрессионного анализа.
27. Теория вероятностей.
28. Случайные события.
29. Пространство элементарных событий.
30. Статистическая вероятность.
31. Условные вероятности.
32. Последовательности независимых испытаний.
33. Случайные величины.
34. Функции и числовые характеристики случайных величин.
35. Основные законы распределения случайных величин.

36. Экспоненциальное распределение и его числовые характеристики.

37. Предельные теоремы теории вероятностей.

**Таблица 9. Примеры оценочных средств с ключами правильных ответов**

№ п/п	Тип задания	Формулировка задания	Правильный Ответ	Время выполнения (в минутах)
1.	Задание закрытого типа	На плоскости нарисованы две концентрические окружности, радиусы которых 6 и 12 см соответственно. Какова вероятность того, что точка брошенная наудачу в большой круг, попадет в кольцо, образованное указанными окружностями? 1) 0,5; 2) 0,65; 3) 0,12; 4) 0,75; 5) 0,60.	4	2
2.		На сборку попадают детали с двух автоматов: 80 % из первого и 20 % из второго. Первый автомат дает 10 % брака, второй – 5 % брака. Найти вероятность попадания на сборку доброкачественной детали. 1) 0,90; 2) 0,09; 3) 0,91; 4) 0,85; 5) 0,15	3	3
3.		Случайная величина распределена по нормальному закону, причем $M(X) = 15$ . Найти $P(10 < X < 15)$ , если известно, что $P(15 < X < 20) = 0,25$ . 1) 0,10; 2) 0,15; 3) 0,20; 4) 0,25; 5) 0,30.	4	3
4.		По выборке объема $n = 51$ найдена смещенная оценка $D_{в=3}$ генеральной дисперсии. Найти несмещенную оценку дисперсии генеральной совокупности. 1) 3,05; 2) 3,06; 3) 3,51; 4) 3,60; 5) 0.	2	3
5.		Как называется число $m_0$ (наступления события в $n$ независимых испытаниях, в каждом из которых вероятность появления события равна $p$ ), определяемое из неравенства: $np - q \leq m_0 \leq np + p$ 1) наибольшее 2) оптимальное	3	3

№ п/п	Тип задания	Формулировка задания	Правильный Ответ	Время выполнения (в минутах)
		3) наивероятнейшее 4) невозможное 5) минимальное		
6.	Задание открытого типа	Задан закон распределения случайной величины X:  Варианта $x_i$ 0    1    2    3 Частота $n_i$ 0,2   0,3   0,4   0,1  Дисперсия $Dx$ равна	0,84	8
7.		Студентам нужно сдать 4 экзамена за 6 дней. Сколькими способами можно составить расписание сдачи экзаменов?	360	8
8.		Задан закон распределения случайной величины X:  Варианта $x_i$ 0    1    2    3 Частота $n_i$ 0,2   0,3   0,4   0,1  Дисперсия $Sx$ равна. Ответ округлить до двух знаков после запятой	0,92	8
9.	Задание комбинированного типа	Задан закон распределения случайной величины X:  Варианта $x_i$ 0    1    2    3 Частота $n_i$ 0,2   0,3   0,4   0,1  Математическое ожидание M: а) 10 б) 6,5 в) 1,4 г) 2,1	Математическое ожидание :  $M = (0*0,2 + 1*0,3 + 2*0,4 + 3*0,1) = 1,4$	8
10.	Задание комбинированного типа	Данное предприятие в среднем выпускает 20 % продукции высшего сорта и 70 % продукции первого сорта. Найти вероятность P того, что случайно взятое изделие этого предприятия будет не высшего или не первого сорта. а) 0,1 б) 0,2 в) 1,1 г) 0,7	Высший сорт = 20%=0,2.  Первый сорт = 70%=0,7.  Вероятность "не высшего и не первого сорта" $P = 1 - (0,2 + 0,7) = 0,1$  Ответ: P = 0,1	8

Полный комплект оценочных материалов по дисциплине (модулю) (фонд оценочных средств) хранится в электронном виде на кафедре, утверждающей рабочую программу дисциплины (модуля), и в Центре мониторинга и аудита качества обучения.

#### 7.4. Методические материалы, определяющие процедуры оценивания результатов обучения по дисциплине (модулю)

**Методические рекомендации по выполнению практических и контрольных работ, проведению зачета**

По дисциплине «Вероятностно-статистические методы в анализе данных» итоговой формой отчетности в 3 семестре является зачет. Согласно действующей в АГУ системе оценивания БАРС для дисциплин, итоговой формой отчетности для которых является зачет, отводится 100 баллов (90 баллов на текущие формы контроля и до 10 баллов отводится на бонусы), которые накапливаются студентом в течение всего семестра изучения дисциплины и распределяются по возможности равномерно по всему семестру. Система накопления баллов, а также система штрафов, представлена в технологической карте дисциплины «Вероятностно-статистические методы в анализе данных».

#### **Отчет по практической работе**

Отчет по практической работе представляется в электронном виде. Защита отчета проходит в форме доклада студента по выполненной работе и ответов на вопросы преподавателя. В случае, если оформление отчета и поведение студента во время защиты соответствуют указанным требованиям, студент получает максимальное количество баллов.

Основаниями для снижения количества баллов в диапазоне от max до min являются:

- небрежное выполнение,
- отсутствие выводов,
- нарушение сроков предоставления отчета.

Отчет не может быть принят и подлежит доработке в случае:

- отсутствия необходимых разделов,
- отсутствия необходимого графического материала,
- неверных результатов расчета.

#### **Критерии оценки по практическим работам:**

– оценка «отлично» выставляется обучающемуся, если студент продемонстрировал глубокие знания теоретического материала и умение их применять, обоснованно изложил свои мысли, сделал необходимые выводы и учел основные нормативно-правовые документы по информационной безопасности;

– оценка «хорошо» выставляется обучающемуся, если студент продемонстрировал глубокие знания теоретического материала и умение их применять, обоснованно изложил свои мысли, сделал необходимые выводы и учел основные нормативно-правовые документы по информационной безопасности, допущены некоторые неточности, имеется одна негрубая ошибка.

– оценка «удовлетворительно» выставляется обучающемуся, если студент ответил на вопросы преимущественно верно, имеются затруднения в формулировке выводов, имеются одна или две негрубые ошибки, учтены не все нормативно-правовые документы по информационной безопасности;

– оценка «неудовлетворительно» выставляется обучающемуся, если студент не дал ответы на поставленные вопросы, обоснования неверные, либо дан верный ответ без его обоснования, сделаны грубые ошибки, отсутствуют знания нормативно-правовых документов по информационной безопасности.

#### **Контрольные работы**

Контрольная работа состоит из 2-х заданий.

Основаниями для снижения оценки за задание являются:

- ошибки в объяснениях и комментариях при верно выполненном задании;
- неполный ответ для теоретических заданий;
- небрежное выполнение;
- многократное переписывание контрольной работы.

Задание не может быть засчитано, если:

- даны два неверных ответа на теоретические вопросы.

#### **Критерии оценки контрольных работ:**

– оценка «отлично» выставляется обучающемуся, если студент продемонстрировал глубокие знания теоретического материала и умение их применять, обоснованно изложил свои мысли, сделал необходимые выводы и учел основные нормативно-правовые документы по информационной безопасности;

– оценка «хорошо» выставляется обучающемуся, если студент продемонстрировал глубокие знания теоретического материала и умение их применять, обоснованно изложил свои мысли, сделал необходимые выводы и учел основные нормативно-правовые документы по информационной безопасности, допущены некоторые неточности, имеется одна негрубая ошибка.

– оценка «удовлетворительно» выставляется обучающемуся, если студент ответил на вопросы преимущественно верно, имеются затруднения в формулировке выводов, имеются одна или две негрубые ошибки, учтены не все нормативно-правовые документы по информационной безопасности;

– оценка «неудовлетворительно» выставляется обучающемуся, если студент не дал ответы на поставленные вопросы, обоснования неверные, либо дан верный ответ без его обоснования, сделаны грубые ошибки, отсутствуют знания нормативно-правовых документов по информационной безопасности.

В соответствии с балльно-рейтинговой системой БАРС по дисциплине отводится 100 баллов (90 баллов на текущие формы контроля и до 10 баллов отводится на бонусы), которые накапливаются студентом в течение всего семестра изучения дисциплины.

Оценивание студентов на **зачете** осуществляется в соответствии с требованиями и критериями 100-балльной шкалы. Учитываются как результаты текущего контроля, так и знания, навыки и умения, непосредственно показанные студентами в ходе зачета.

Текущий контроль осуществляется в ходе учебного процесса и консультирования студентов, по результатам выполнения лабораторных и контрольной работ. Он предусматривает проверку готовности студентов к плановым занятиям, оценку качества и самостоятельности выполнения заданий на практических занятиях, проверку правильности решения задач, выданных на самостоятельную проработку.

На зачете осуществляется комплексная проверка знаний, навыков и умений студентов по всему теоретическому материалу дисциплины и с проверкой практических навыков и умений по разработке документов различных видов. Теоретические знания оцениваются путем компьютерного тестирования или на основании письменных ответов студентов по нескольким теоретическим вопросам.

**Таблица 10. Технологическая карта рейтинговых баллов по дисциплине (модулю)**

№ п/п	Контролируемые мероприятия	Количество мероприятий / баллы	Максимальное количество баллов	Срок представления
<b>Основной блок</b>				
– 1	<i>Выполнение лабораторной работы</i>	5/15	75	По расписанию
– 2	<i>Выполнение контрольной работы</i>	1/10	10	
– 3	<i>Вопросы к зачету</i>	1/5	5	
<b>Всего</b>			<b>90</b>	-
<b>Блок бонусов</b>				
– 4	<i>Посещение занятий без пропусков</i>	1	3	
– 5	<i>Своевременное выполнение всех заданий</i>	1	3	
– 6	<i>Активность студента на</i>	1	4	

№ п/п	Контролируемые мероприятия	Количество мероприятий / баллы	Максимальное количество баллов	Срок представления
	<i>занятия</i>			
<b>Всего</b>			<b>10</b>	-
<b>ИТОГО</b>			<b>100</b>	-

**Таблица 11. Система штрафов (для одного занятия)**

Показатель	Балл
<i>Опоздание на занятие</i>	- 1
<i>Нарушение учебной дисциплины</i>	- 1
<i>Неготовность к занятию</i>	- 2
<i>Пропуск занятия без уважительной причины</i>	- 2

**Таблица 12. Шкала перевода рейтинговых баллов в итоговую оценку за семестр по дисциплине (модулю)**

Сумма баллов	Оценка по 4-балльной шкале	
90–100	5 (отлично)	зачтено
85–89	4 (хорошо)	
75–84		
70–74		
65–69	3 (удовлетворительно)	
60–64	2 (неудовлетворительно)	Не зачтено
Ниже 60		

При реализации дисциплины (модуля) в зависимости от уровня подготовленности обучающихся могут быть использованы иные формы, методы контроля и оценочные средства, исходя из конкретной ситуации.

## **8. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)**

### **8.1. Основная литература:**

1. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2020. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/450262>
2. Теория вероятностей и математическая статистика. Практикум [Электронный ресурс]: учеб. пособие / И.Ю. Мацкевич, Н.П. Петрова, Л.И. Тарусина - Минск : РИПО, 2017. - URL: <http://www.studentlibrary.ru/book/ISBN9789855037119.html> (ЭБС «Консультант студента»).
3. Ракитский, А. А. Методы машинного обучения : учебно-методическое пособие / А. А. Ракитский. — Новосибирск : Сибирский государственный университет телекоммуникаций и информатики, 2018. — 32 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/90591> .
4. Рашка С., Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения / Рашка С. - М. : ДМК Пресс, 2017. - 418 с. - ISBN 978-

5-97060-409-0 - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : <https://www.studentlibrary.ru/book/ISBN9785970604090.html>

5. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2020. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/450166>

## **8.2. Дополнительная литература:**

1. Теория вероятностей. Примеры и задачи [Электронный ресурс]: учеб. пособие / Васильчик М.Ю. - Новосибирск : Изд-во НГТУ, 2014. - URL: <http://www.studentlibrary.ru/book/ISBN9785778224872.html> (ЭБС «Консультант студента»).

2. Теория вероятностей и математическая статистика [Электронный ресурс] / Балдин К. В. - М. : Дашков и К, 2014. - URL: <http://www.studentlibrary.ru/book/ISBN9785394021084.html> (ЭБС «Консультант студента»).

3. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. — Москва : Московский городской педагогический университет, 2012. — 204 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/26444.html>

4. Билл, Фрэнкс Укрощение больших данных : как извлекать знания из массивов информации с помощью глубокой аналитики / Фрэнкс Билл ; перевод А. Баранов. — Москва : Манн, Иванов и Фербер, 2014. — 340 с. — ISBN 978-5-00057-146-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/39433.html>

5. Флах П., Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / Флах П. - М. : ДМК Пресс, 2015. - 400 с. - ISBN 978-5-97060-273-7 - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : <https://www.studentlibrary.ru/book/ISBN9785970602737.html>

## **8.3. Интернет-ресурсы, необходимые для освоения дисциплины (модуля)**

1. Электронно-библиотечная система (ЭБС) ООО «Политехресурс» «Консультант студента». Многопрофильный образовательный ресурс «Консультант студента» является электронной библиотечной системой, предоставляющей доступ через сеть Интернет к учебной литературе и дополнительным материалам, приобретенным на основании прямых договоров с правообладателями. Каталог в настоящее время содержит около 15000 наименований. [www.studentlibrary.ru](http://www.studentlibrary.ru).

2. Электронная библиотека «Астраханский государственный университет» собственной генерации на платформе ЭБС «Электронный Читальный зал – БиблиоТех». <https://biblio.asu.edu.ru>

## **9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)**

Для проведения лекционных занятий необходима мультимедийная аудитория, оснащенная компьютерной презентационной техникой.

Для проведения публичной защиты проектов, необходима мультимедийная аудитория с проектором.

Для проведения лабораторных занятий необходима компьютерная аудитория, в которой организован доступ к сети Интернет и установлено программное обеспечение:

## **10. ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ (МОДУЛЯ) ПРИ ОБУЧЕНИИ ИНВАЛИДОВ И ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ**

Рабочая программа дисциплины (модуля) при необходимости может быть адаптирована для обучения (в том числе с применением дистанционных образовательных технологий) лиц с ограниченными возможностями здоровья, инвалидов. Для этого требуется заявление обучающихся, являющихся лицами с ограниченными возможностями здоровья, инвалидами, или их законных представителей и рекомендации психолого-медико-педагогической комиссии. При обучении лиц с ограниченными возможностями здоровья учитываются их индивидуальные психофизические особенности. Обучение инвалидов осуществляется также в соответствии с индивидуальной программой реабилитации инвалида (при наличии).

Для лиц с нарушением слуха возможно предоставление учебной информации в визуальной форме (краткий конспект лекций; тексты заданий, напечатанные увеличенным шрифтом), на аудиторных занятиях допускается присутствие ассистента, а также сурдопереводчиков и тифлосурдопереводчиков. Текущий контроль успеваемости осуществляется в письменной форме: обучающийся письменно отвечает на вопросы, письменно выполняет практические задания. Доклад (реферат) также может быть представлен в письменной форме, при этом требования к содержанию остаются теми же, а требования к качеству изложения материала (понятность, качество речи, взаимодействие с аудиторией и т. д.) заменяются на соответствующие требования, предъявляемые к письменным работам (качество оформления текста и списка литературы, грамотность, наличие иллюстрационных материалов и т.д.). Промежуточная аттестация для лиц с нарушениями слуха проводится в письменной форме, при этом используются общие критерии оценивания. При необходимости время подготовки к ответу может быть увеличено.

Для лиц с нарушением зрения допускается аудиальное предоставление информации, а также использование на аудиторных занятиях звукозаписывающих устройств (диктофонов и т.д.). Допускается присутствие на занятиях ассистента (помощника), оказывающего обучающимся необходимую техническую помощь. Текущий контроль успеваемости осуществляется в устной форме. При проведении промежуточной аттестации для лиц с нарушением зрения тестирование может быть заменено на устное собеседование по вопросам.

Для лиц с ограниченными возможностями здоровья, имеющих нарушения опорно-двигательного аппарата, на аудиторных занятиях, а также при проведении процедур текущего контроля успеваемости и промежуточной аттестации могут быть предоставлены необходимые технические средства (персональный компьютер, ноутбук или другой гаджет); допускается присутствие ассистента (ассистентов), оказывающего обучающимся необходимую техническую помощь (занять рабочее место, передвигаться по аудитории, прочитать задание, оформить ответ, общаться с преподавателем).